Localization of the RNA-binding domain of mouse hepatitis virus nucleocapsid protein

Gary W. Nelson1 and Stephen A. Stohlman1,2*

1Department of Microbiology and 2Department of Neurology, University of Southern California
School of Medicine, Los Angeles, California 90033, U.S.A.

The 454-amino acid nucleocapsid (N) protein of mouse hepatitis virus (MHV) binds the leader RNA sequence located at the 5' ends of all plus-sense genomic and subgenomic viral mRNAs. Purified N protein was cleaved with formic acid to determine which domain interacts with the leader RNA sequence. Incubation at 42 °C resulted in partial cleavage into two fragments of M₁ of approximately 32K and 37K and three fragments of 17K, 16K and 14K. Incubation at 56 °C resulted in complete cleavage yielding only the three lower molecular mass products. Both the 32K and 37K partial cleavage products and one of the complete cleavage products bind MHV leader RNA, suggesting that the central region of the N protein contains the RNA-binding domain. Monoclonal antibody mapping of the cleavage products confirmed that the MHV leader RNA binding domain is contained within the central 140-amino acid fragment, comprising amino acids 169 to 308. Analysis of the amino acids within this domain indicates no similarity to any previously described RNA-binding protein, suggesting that N protein may possess a unique RNA-binding motif.

The murine coronavirus, mouse hepatitis virus (MHV), is an enveloped, single-stranded RNA virus containing a positive-polarity genome of approximately 31 kb (Lee et al., 1990). The virion is composed of three or four structural proteins (Lai, 1990), the spike, membrane and the optional haemagglutinin–esterase proteins, all of which are integral membrane proteins. The fourth protein, the nucleocapsid (N) protein, is a basic phosphoprotein with a pI of approximately 10.4. It is phosphorylated exclusively on serine residues (Stohlman & Lai, 1979) and interacts with genomic RNA to form the helical nucleocapsid structure (Sturman et al., 1980). Comparison of the N protein coding sequences of five MHV strains has revealed that the proteins vary from 454 to 455 amino acids in length and are more than 90% identical (Parker & Masters, 1990).

Analysis of the amino acid sequence of the A59 N protein (Armstrong et al., 1983; Parker & Masters, 1990) indicated that formic acid, which hydrolyses peptide bonds between aspartic and proline residues and has been used to analyse a variety of proteins (Landon, 1977), would cleave the N protein into three fragments: an amino-terminal fragment of 168 amino acids (19K; designated A), a middle fragment of 140 amino acids (15K; designated B) and a carboxy-terminal fragment of 146 amino acids (16K; designated C) (Fig. 1a).

N protein was isolated from monolayers of DBT cells infected with the A59 strain of MHV for 1 h at room temperature at an m.o.i. of 5 to 10. At 5–6 h post-infection (p.i.) the cultures were placed on ice, washed once with NTE (100 mM-NaCl, 10 mM-Tris–HCl pH 7.4, 1 mM-EDTA) and lysed using a buffer containing 150 mM-NaCl, 100 mM-Tris–HCl pH 7.4, 0.5% NP40, 0.02% PMSF (Sigma). Nuclei and large debris were removed by centrifugation at 1500 g for 5 min. Supernatants were mixed with an equal volume of 2 × sample buffer (LSB) (Laemmli, 1970), boiled for 3 min and stored at -20 °C. Lysates were separated on 20 cm 10% SDS-polyacrylamide gels using lysates of 35S-labelled A59-infected cells as markers. N protein was recovered by diffusion into three gel volumes of 10 mM-Tris–HCl pH 8.8 with gentle agitation at 4 °C for 18 h, lyophilized and resuspended in a small volume of water. The SDS and Tris–HCl/glycine concentrations were reduced by electroelution in 12 mM-Tris–HCl pH 8.3 with 96 mM-glycine for 18 h at 90 V constant voltage (Schleicher & Schuell). Protein was quantified by the Bio-Rad protein assay (Bio-Rad Laboratories). Approximately 10 μg lyophilized N protein was resuspended in water and three volumes of 99% formic acid (Sigma) were added. Samples were incubated at 42 °C or 56 °C for 18 h, lyophilized and resuspended in a small volume of water. The SDS and Tris–HCl/glycine concentrations were reduced by electroelution in 12 mM-Tris–HCl pH 8.3 with 96 mM-glycine for 18 h at 90 V constant voltage (Schleicher & Schuell). Protein was quantified by the Bio-Rad protein assay (Bio-Rad Laboratories). Approximately 10 μg lyophilized N protein was resuspended in water and three volumes of 99% formic acid (Sigma) were added. Samples were incubated at 42 °C or 56 °C for 18 h, lyophilized and resuspended in LSB. The pH was adjusted with 5 M-NaOH until the dye remained blue. Cleavage was analysed by SDS–PAGE following treatment at either 42 °C or 56 °C. Treatment at 42 °C for 18 h resulted in partial cleavage yielding two inter-
mediates, designated I and II, with apparent Mr values of 37K (I) and 32K (II) in addition to the three predicted complete cleavage products, designated P1, P2 and P3, with apparent Mr values of 17K, 16K and 14K respectively (Fig. 1b). Treatment at 56 °C caused complete cleavage (Fig. 1b). The B fragment contains four potential 35S-labeling sites compared to only a single site for the other two fragments (Fig. 1a). Densitometric analysis of the autoradiogram yields a 1:3:2:1 ratio (P1: P2: P3), suggesting that P2 represents the B fragment.

Monoclonal antibodies (MAbs) specific for the N protein (designated A. 1.10, A. 2.16, A. 2.17, A. 2.6, A. 3.1 and A. 3.7) were used to verify the identity of the formic acid-derived peptides (Gilmore et al., 1987). Cleavage fragments were separated by 13.5% SDS-PAGE (Laemmli, 1970) and transferred to nitrocellulose membranes (Bio-Rad) as previously described (Stohlman et al., 1992). Fig. 2 shows that none of the N-specific MAbs recognize the P3 fragment (see below). MAb A.2.6 recognized both intermediate fragments and the P2 fragment. The epitope recognized by MAb A.2.6 resides within the B region, as determined by ELISA using truncated N proteins of the JHM strain of MHV (JHMV) expressed by recombinant vaccinia viruses (Stohlman et al., 1992). MAb A.3.1, which also binds an epitope within the B region, exhibited a similar pattern of reactivity (data not shown). These data confirm that the P2 fragment is derived from the B region of the N protein since MAb A.2.6 recognizes both intermediates and the P2 cleavage product, consistent with the assumption that the intermediates arise from loss of either the carboxy- or amino-terminal fragments.

MAb A.2.16 recognizes intermediate II and the P1 fragment (Fig. 2). Similar results were also obtained with MAb A.1.10 (data not shown). The derivation of P1 from the carboxy end of the N protein has been confirmed by ELISA using MAb A.2.16 and the recombinant vaccinia virus-expressed N protein truncations as antigens, thus identifying intermediate II and P1 as B/C and C respectively.

Fig. 1. Formic acid-derived cleavage products of the N protein. (a) Schematic diagram of the predicted complete cleavage pattern. Cleavage sites are between amino acids 168/169 and 308/309 at adjacent aspartic acid (D) and proline (P) residues. The predicted Mr's of the three peptides (A, B, C) are indicated. The position of 35S-radiolabelled methionine (M) and cysteine (C) residues radiolabelled with trans-35S are also identified. (b) Trans-35S-labelled untreated N protein (lane 1), treated with formic acid at 42 °C (lane 2) or 56 °C (lane 3) was separated by 13.5% SDS-PAGE.

Fig. 2. Analysis of the N protein cleavage products by MAb binding. A mixture of untreated, 42 °C or 56 °C formic acid-treated N protein was tested with N-specific MAbs by Western blotting. Formic acid cleavage intermediates are labelled I (A/B) and II (B/C) and the two final cleavage products recognized are labelled P1 (C) and P2 (B). Lane 1, a mixture of eight MAbs specific for the N protein; lane 2, MAb A.2.16; lane 3, MAb A.2.17; lane 4, MAb A.2.6.
MAb A. 2.17 recognized intermediate I but none of the final cleavage products (Fig. 2), suggesting that this MAb may recognize a conformational epitope or an epitope adjacent or mapping at the A/B cleavage site. Similar results were obtained using MAb A. 3.7 (data not shown). The epitope(s) recognized by MAbs A. 2.17 and A. 3.7 could not be identified by ELISA since neither exhibits reactivity to JHMV (Gilmore et al., 1987). In addition, no immunoreactivity to a recombinant vaccinia virus-expressed amino-terminal peptide (amino acids 1 to 134) of the JHMV N protein has been detected using immune sera derived from BALB/c or C57BL/6 mice (Stohlman et al., 1992), suggesting that MAbs A. 2.17 and A. 3.7 recognize a conformational epitope(s) and that the P3 fragment is indeed the A fragment and is derived from the amino terminus of the N protein.

The results shown in Fig. 1(b) and 2 indicate that the calculated \( M_r \) of the A (P3) fragment does not correspond to the apparent \( M_r \) following formic acid cleavage (Fig. 1). From the expected pattern (Fig. 1a) fragment A would be the largest (19K). The data, however, indicate that the A fragment migrates at an apparent \( M_r \) of 14K whereas the B and C fragments migrate at approximately their expected positions. Treatment with formic acid for fewer than 18 h results in a faint band at the predicted \( M_r \) of the A fragment (19K). After 18 h incubation, however, this band is barely detectable and the lower (P3) band predominates. These data suggest that the A fragment is cleaved at a secondary site or exhibits anomalous electrophoretic mobility, resulting in reduction of the apparent \( M_r \). Alternatively, the P3 fragment could represent a cleavage product derived from another region of the protein.

To determine whether the cleavage products retained RNA-binding activity, unlabelled and \(^{35}S\)-labelled N proteins were treated with formic acid at either 42 °C or 56 °C. Unlabelled or \(^{35}S\)-labelled cleavage products were separated in adjacent lanes by SDS–PAGE. Figure 3(a) shows the \(^{35}S\)-labelled products. To determine whether the cleavage products retained RNA-binding activity the membrane was incubated with a \(^{32}P\)-labelled riboprobe containing the MHV leader RNA sequence (Stohlman et al., 1988). Fig. 3(b) shows that both intermediates contain the RNA-binding site. However, among the complete cleavage products only the B fragment binds RNA. These data confirm that the two partial cleavage products result from the loss of either the C region (fragment I) or the A region (fragment II) and indicate that the B fragment contains the RNA-binding domain.

N protein fragments obtained after complete cleavage were also tested with an MHV leader RNA riboprobe in the presence of competitor RNA at levels previously shown to eliminate non-specific interactions (Stohlman et al., 1988). The B fragment was the only cleavage product that bound RNA (Fig. 4a). To evaluate whether MHV leader RNA binding of the B fragment was equivalent to that of the native N protein, untreated and N proteins treated with formic acid at 42 °C or 56 °C were mixed at equimolar ratios, separated by SDS–PAGE, transferred to nitrocellulose membranes, and tested with the MHV leader \(^{32}P\)-labelled riboprobe in the presence of competitor RNA. The intermediate cleavage products and the B fragment exhibited RNA-binding properties similar to the native N protein (Fig. 4b), confirming that the RNA-binding domain resides within this region of the N protein.

Based on a comparison of the amino acid sequences of a number of MHV strains, the highly conserved centre of the molecule was predicted to contain a potential RNA-binding domain (Parker & Masters, 1990). Recently, the analysis of truncated N proteins expressed by \textit{in vitro} translation demonstrated that the central domain exhibits RNA-binding activity (Masters, 1992). However, in contrast to our previous results (Stohlman et al., 1988) and the data contained in this manuscript, no specificity for the MHV leader RNA could be detected (Masters, 1992). The reason(s) for the discrepancy between these data is unclear. However, the difference may be due to the two sources of N protein used, i.e. N protein prepared by \textit{in vitro} translation (Masters, 1992) as opposed to isolated native protein. N protein is phosphorylated exclusively on serine residues (Stohlman et al., 1992), suggesting that MAbs A.2.17 and A. 3.7 could not be identified by ELISA since neither exhibits reactivity to JHMV.
protein, intermediates and final cleavage products were separated by polyacrylamide gel electrophoresis and identified at the top of each lane. (b) An equimolar mixture of N protein treated at 56 °C for 18 h, separated by 13.5% SDS–PAGE and transferred to nitrocellulose was exposed to radiolabelled MHV leader RNA in the presence of competitor RNA at concentrations (μg/ml) as indicated.

Fig. 4. North-western blot of N protein cleavage products. (a) N protein treated at 56 °C for 18 h, separated by 13.5% SDS–PAGE and transferred to nitrocellulose was exposed to radiolabelled MHV leader RNA in the presence of competitor RNA at concentrations (μg/ml) identified at the top of each lane. (b) An equimolar mixture of N protein, intermediates and final cleavage products were separated by 13.5% SDS–PAGE, transferred to nitrocellulose and exposed to radiolabelled MHV leader RNA in the presence of competitor RNA (μg/ml) as indicated.

and Lai, 1979), and the B peptide contains 16 serine residues as potential phosphorylation sites. Formic acid cleavage of 32P-labelled N protein prepared from infected cells indicates that the B fragment is indeed phosphorylated (data not shown). If phosphorylation plays a role in the RNA-binding activity of N protein, as has been suggested previously (Mohandas & Dales, 1991), it is possible that the absence of appropriate phosphorylation could have altered RNA-binding characteristics.

The precise details of MHV mRNA transcription are unclear. Antibody specific for the N protein blocks viral mRNA synthesis in vitro (Compton et al., 1987) implying that the N protein plays a role in transcription. The leader RNA, which acts as a primer for mRNA synthesis (Lai, 1990), is transcribed independently from the 3′ end of genomic-length negative-strand templates (Baric et al., 1985). The 3′ region of leader RNA contains two or three repeats of a consensus sequence (UCUAA) complementary to negative-strand intergenic regions (Budzilowicz et al., 1985; Sheih et al., 1987). During transcription, leader RNA sequences are cleaved and fused to mRNA bodies at a junction site within the second or third (3′) tandem repeat of the consensus sequence (Baker & Lai, 1990; Joo & Makino, 1992). The N protein forms high affinity complexes with the leader RNA through interaction at the 3′ end of leader RNA which contains the consensus repeat (Baric et al., 1988; S. A. Stohlman et al., unpublished). Therefore, N protein may function to target the transcription complex to the appropriate subcellular sites of transcription or interact with the viral polymerase to direct the site of leader–message fusion by protecting bound sequences from polymerase nuclease activity (Sheih et al., 1987; Joo & Makino, 1992). In addition, the binding of N protein to leader RNA may prevent annealing of newly transcribed RNA or the protein may act as a helicase unwinding newly transcribed RNA from the negative-strand templates. Recent evidence suggests that N protein bound to MHV mRNAs may also serve to enhance translation in infected cells (unpublished).

The data in this report indicate that the binding of MHV leader RNA to the N protein is mediated by a central domain between amino acids 169 and 308. As with other nucleic acid-binding proteins (Weeks et al., 1990; Scherly et al., 1990; Lutz-Freyermuth et al., 1990; Adam et al., 1986; Sachs et al., 1987), the isolated RNA-binding domain of the N protein retains the ability to interact specifically with its target RNA sequence. The amino acids within this central domain were analysed by sequence comparison and demonstrate no similarity to other known nucleic acid-binding proteins or binding motifs. This N protein domain contains multiple basic residues similar to other RNA-binding proteins; however, a general electrostatic interaction of the B domain with RNA is unlikely to be solely responsible for binding to leader RNA because both the amino-terminal and middle fragments of the N protein have pIs greater than 11. The N proteins in four of the five sequenced MHV strains are 100% identical within the B domain (Parker & Masters, 1990), indicating that this domain performs a critical biological function.

Amino acid sequence analysis of the N proteins from a variety of coronavirus strains suggests that a highly conserved domain within the amino-terminal portion of this central domain may contain the minimum RNA-binding domain. Experiments to test this possibility directly are currently in progress. However, this domain does not contain an RNA-binding motif similar to any previously identified RNA-binding motif; therefore, the elucidation of which amino acids are responsible for binding MHV leader RNA should contribute to the understanding of MHV replication as well as the elements required for RNA–protein interactions.

We thank E. Dimacali and L. Pen for technical assistance, Sonia Q. Garcia for preparation of this manuscript, and Wendy Gilmore, Stanley Tahara and John Polo for helpful discussions. This work was supported by NIH grant AI-17447.
supported by the National Science Foundation grant DMB17148 and the National Institute of Health grants AI07078 and NS18146.

References


(Received 30 October 1992: Accepted 13 April 1993)